Where to Start When Designing a Research Project: Part I

Mario Davidson, PhD
Department of Biostatistics
Vanderbilt University School of Medicine
Disclosure of Financial Interests

- No Financial Conflicts
- No Commercial Interests
- No Spousal Conflicts
Introduction

- Bird's Eye View Workshop
- Created for Beginners; Applicable to All
- Model Components on Slides
- Intermediate and Advanced Should Frame Into Proposal
Introduction

- Part I
  - The Research Question
  - Feasibility
  - Qualitative Research
  - Mixed Methods

- Part II
  - Quantitative Research
  - Measurements and Outcomes
  - Getting Statistical Significance
Formulating a Research Question

Objectives

At the end of this session, you will:

- Know the major components needed to design a research project
- Know important questions and issues to consider when creating a research project
- Be able to list four feasibility concerns that can change the research question
Research Proposal Template

STUDY TITLE

COMMITTEE
Your name, mentor, and others working on the project.

RESEARCH QUESTION
Concise, informative question that outlines and guides the methods, population, and other details pertinent to developing the study.

RATIONALE
Include relevant background information, references, and potential implications of results.

STUDY DESIGN
Type of study proposed (e.g. randomized controlled trial, prospective cohort study, etc.). Explains the way data will be accessed and gathered.

NATURE OF INTERVENTION(S)
Description of study products, counseling programs, etc. (if applicable).

STUDY POPULATIONS
Salient characteristics of target study population (key inclusion/exclusion criteria).

STUDY SIZE OR POWER
An estimate of total number of participants needed in the study and number in each arm, if applicable or the power to detect an effect. Request biostatistics help if necessary.

HYPOTHESIS(Optional)
An assumption taken to be true for the purpose of argument or investigation. Expected outcome(s) of the study.

FEASIBILITY
Ability to answer question fully, resources, barriers, timeline, supplies, scheduling, data collection, analyses, etc.

DATA(Optional)
Type of data needed to be collected.

INSTRUMENTS(Optional)
Tools needed to measure data.
In qualitative research, some have added a separate “Ethical” component to the model.
The Qualitative Planning Stage: Goals

- Is the study worth doing?
- What issues do you want to clarify?
- Why should anyone care about the results?
The Qualitative Planning Stage: Conceptual Framework

- What is going on with the issues, settings, or people you plan to study?
- What beliefs, theories, ideas, and/or finding will you draw upon to inform the research and what literature, experiences, and preliminary studies will you draw on for understanding the phenomenon?
- Conceptual Context of Theory and Literature
  - Will Theory Change the Lens?
The Qualitative Planning Stage: Methods

- What will you do in conducting the study?
- What techniques will you use to collect and analyse your data?

**Sampling Design**
- Site Selection
- Gaining Access

**Analysis**
- Coding
- Software
- Memo Writing

**Data Collection Techniques**
- Interviews
- Surveys
- Observation
- Focus Groups
- Print Media
- Chart Review
The Qualitative Planning Stage: Validity

- How might the findings be incorrect?
- Are there plausible interpretations and validity threats? If so, how will you deal with these?
- How can the collected data challenge your ideas?
- Why should anyone believe your results?
Qualitative Research: Validity

- Triangulation
- Prolonged Engagement
- Rich, Thick Description
- Peer Review
- Negative Case Analysis
- Researcher Bias
- Member Checking
The Qualitative Planning Stage: Research Question

- What do you specifically want to learn about?
- What do you not know about that you would like to learn?
- What questions will you try to learn about and their relation to other questions?
The Qualitative Planning Stage: Research Question

- We will focus on the research question.
- Maxwell says the question should be responsive to the other components.
- In quantitative, the research question dictates the other components.
The Quantitative Planning Phase

- Components of Model on Each Slide

- State questions and hypotheses, identify variables
- Determine design structure
- Identify population and sample
- Design instruments and classify: operational definitions
- Select statistical test for resolving hypotheses

(Black, 1999; Doing Quantitative Research)
The research question dictates:

- The goal of the study
- A summary of what is to be achieved
- The methodology
- The who, what, when, why, how, where, condition, intervention questions
- The type of analysis

If you find yourself trying to gather information not tied to your question, leave it alone or revise your question appropriately!
The Research Question: Types

- Types of Questions
  - Descriptive (What are the number of hours worked by residents?)
  - Exploratory (How do VMS III's deal with end of life patients?)
  - Explanatory – Explain relations/patterns (Are MSTP students more engaged in reading articles than medical students?)
  - Predictive - Does the workload of getting a MD and/or PhD predict couples' ability to stay together?

- Key Words: Improvement, Perceive, Difference, Association, Trends, Describe, View, etc.
- Use Assessment Questions (Primary Sources, Relevant Authors, Search Complete, Methodology Good)
- Complete When References Repeated
Group Activity: Formulate a Question

- Topic: The cost of health care
- Budget: $2,000
- Time frame: 2 years
- Take one minute to brainstorm and write words associated with the topic
Group Activity: Formulate a Question

- Topic: The cost of health care
- Write down as many research questions using brainstorming words (or additional words) related to the topic
- Choose the best question
Discuss the following questions below. Are they researchable? Good questions? Too broad? Too narrow?

- Is the rising cost of health care too much for the average American?
- What is the average copay of Dr. Garrett's patients?
Characteristics of a Good Research Question

- Does the question answer the who, what, when, why, how, where, condition, and/or intervention?
- FINER (Hulley, Cummings, et. al)
  - Feasible
  - Interesting
  - Novel
  - Ethical
  - Relevant
Feasibility

- Cost
  - Staff
  - Supplies
  - Compensation
  - Continuing Education
  - Travel

- Time
  - Planning
    - Create a timeline!
  - Scheduling
  - Literature Search
  - Data Collection
  - Analyzing
  - Representation
Formulate a Question

- Does the question truly spark your interest?
- What information is required to answer the question?
- Who will you need to research?
- Will you be able to answer the question fully?
- Is there enough information about the question?
- What resources will be required?
- Will the question be able to be answered ethically?
- Will anything need to be measured and how can it be measured?
- What is the relevance?
During their third year of medical school, how do Vanderbilt medical students describe their empathy towards chronically ill patients who express themselves as being in a lot of pain?

How would we design this study (Time permitted)?
Formulate a Question

- Use the Group exercise as a tool
- Concise
- Accurate
- Revise, Revise, Revise Until It's RIGHT!
At the end of this session, you will be able to:

- Understand the strengths and weaknesses of qualitative, quantitative, and mixed methodology
- Define at least 5 reasons why papers are rejected
- Define at least 5 key steps in designing a project
Qualitative, Quantitative, & Mixed Methodology

- **Research Question Dictates Method!**

<table>
<thead>
<tr>
<th>TABLE 2.1</th>
<th>Emphases of Quantitative, Mixed, and Qualitative Research</th>
</tr>
</thead>
</table>
| **Scientific method** | Quantitative Research: Deductive or “top-down” 
The researcher tests hypotheses and theory with data | Mixed Research: Deductive and inductive | Qualitative Research: Inductive or “bottom-up” 
The researcher generates new hypotheses and grounded theory from data collected during fieldwork |
| **View of human behavior** | Quantitative Research: Behavior is regular and predictable | Mixed Research: Behavior is somewhat predictable | Qualitative Research: Behavior is fluid, dynamic, situational, social, contextual, and personal |
| **Most common research objectives** | Quantitative Research: Description, explanation, and prediction | Mixed Research: Multiple objectives | Qualitative Research: Description, exploration, and discovery |
| **Focus** | Quantitative Research: Narrow-angle lens, testing specific hypotheses | Mixed Research: Multilens focus | Qualitative Research: Wide-angle and “deep-angle” lens, examining the breadth and depth of phenomena to learn more about them |
| **Nature of observation** | Quantitative Research: Attempt to study behavior under controlled conditions | Mixed Research: Study behavior in more than one context or condition | Qualitative Research: Study behavior in natural environments |
| **Nature of reality** | Quantitative Research: Objective (different observers agree on what is observed) | Mixed Research: Common sense realism and pragmatic view of world (i.e., what works is what is “real” or true) | Qualitative Research: Subjective, personal, and socially constructed |
| **Form of data collected** | Quantitative Research: Collect quantitative data based on precise measurement using structured and validated data collection instruments (e.g., closed-ended items, rating scales, behavioral responses) | Mixed Research: Multiple forms | Qualitative Research: Collect qualitative data (e.g., in-depth interviews, participant observation, field notes, and open-ended questions) The researcher is the primary data collection instrument |
| **Nature of data** | Quantitative Research: Variables | Mixed Research: Mixture of variables, words, and images | Qualitative Research: Words, images, categories |
| **Data analysis** | Quantitative Research: Identify statistical relationships | Mixed Research: Quantitative and qualitative | Qualitative Research: Search for patterns, themes, and holistic features |
| **Results** | Quantitative Research: Generalizable findings | Mixed Research: Corroborated findings may generalize | Qualitative Research: Particularistic findings |
| **Form of final report** | Quantitative Research: Statistical report (e.g., with correlations, comparisons of means, and reporting of statistical significance of findings) | Mixed Research: Eclectic and pragmatic | Qualitative Research: Narrative report with contextual description and direct quotations from research participants |
Key Steps of Designing a Project

- This is a not an exhaustive, linear process.
- **Throughout the entire process**, talk to key people: mentor(s), biostatistician, IRB representative.
- Create a good research question.
- Review the literature for creativity, inspiration, methodology, design, limitations, gaps, etc.
- Determine the form of data to be collected.
- Consider key theories.
Key Steps of Designing a Project

- Know the type of analysis that will be conducted BEFORE finalizing your research proposal.
- Create a timeline.
- Assure access to the subjects.
- Consider geographical barriers.
- Consider relationship or rapport with participants.
- Be careful of ethical concerns.
- Decide the best way to represent the data.
Reasons Why Papers are Rejected

- Poor research design
- Poor methods section
- Unsupported conclusions
- Unoriginal research
- Poor attention to validation, trustworthiness, reliability
- Failure to collect key variables especially confounders

- Sample not representative of the population
- Sample size too small
- Incorrect analysis
- Author's hypothesis not tested
- Poor writing
STUDY TITLE

COMMITTEE
Your name, mentor, and others working on the project.

RESEARCH QUESTION
Concise, informative question that outlines and guides the methods, population, and other details pertinent to developing the study.

RATIONALE
Include relevant background information, references, and potential implications of results.

STUDY DESIGN
Type of study proposed (e.g. randomized controlled trial, prospective cohort study, etc.). Explains the way data will be accessed and gathered.

NATURE OF INTERVENTION(S)
Description of study products, counseling programs, etc. (if applicable).

STUDY POPULATIONS
Salient characteristics of target study population (key inclusion/exclusion criteria).

STUDY SIZE OR POWER
An estimate of total number of participants needed in the study and number in each arm, if applicable or the power to detect an effect. Request biostatistics help if necessary.

HYPOTHESIS(OPTIONAL)
An assumption taken to be true for the purpose of argument or investigation. Expected outcome(s) of the study.

FEASIBILITY
Ability to answer question fully, resources, barriers, timeline, supplies, scheduling, data collection, analyses, etc.

DATA(OPTIONAL)
Type of data needed to be collected.

INSTRUMENTS(OPTIONAL)
Tools needed to measure data.
Some Available Vanderbilt University Resources

- Daily Biostatistics Clinic  (http://biostat.mc.vanderbilt.edu/wiki/Main/Clinics)
- Biostatistics Collaboration Center  (http://biostat.mc.vanderbilt.edu/wiki/Main/BCC)
- Biostatistics Collaboration Plan  (http://biostat.mc.vanderbilt.edu/wiki/Main/CollaborationAssignments)
- CORE Conversations  (https://medschool.vanderbilt.edu/otlm/core/core_conversations)
- CORE Collaborative in Qualitative Research (limited); Contact Regina Russell at regina.russell@Vanderbilt.Edu
- Design Studio and Vanderbilt Institute for Clinical and Translational Research  (http://www.mc.vanderbilt.edu/victr/pub/)
Part II Foreshadow

- Inclusion and Exclusion Criteria
- Bias
- Getting Significant Results
- Minimizing Sample Size and Increasing Power
- Measures, Outcomes, and Confounding
- Additional Reasons Why Papers are Rejected
Acknowledgments and References

Acknowledgements to Cathy Jenkins, Regina Russell, and Katherine Allen for constructive criticism, editing, and ideas.

Where to Start When Designing a Research Project: Part II

Mario Davidson, PhD
Department of Biostatistics
Vanderbilt University School of Medicine
Disclosure of Financial Interests

- No Financial Conflicts
- No Commercial Interests
- No Spousal Conflicts
Introduction

- Bird's Eye View Workshop
- Created for Beginners; Applicable to All
- Model Components on Slides
- Intermediate and Advanced Should Frame Into Proposal
Recap from Part I

- Maxwell Qualitative Model and Black Quantitative Model
- Research Question – FINER: Feasible, Interesting, Novel, Ethical, Relevant
- Question Dictates Method
- Designing a Project Not Linear
- Literature Review
- Validity in Qualitative Research
In qualitative research, some have added a separate “Ethical” component to the model.

Quantitative Research: Objectives

At the end of the session, you will be able to:

- List the major components needed to design a quantitative research project
- Understand how a biostatistician can help your research
- Define 2 ways of minimizing sample size or maximizing power
- Understand why having the largest sample possible is not necessarily good
The Planning Stage

State questions and hypotheses, identify variables → Determine design structure → Identify population and sample → Design instruments and classify: operational definitions → Select statistical test for resolving hypotheses

(Black, 1999; Doing Quantitative Research)
STUDY TITLE

COMMITTEE
Your name, mentor, and others working on the project.

RESEARCH QUESTION
Concise, informative question that outlines and guides the methods, population, and other details pertinent to developing the study.

RATIONALE
Include relevant background information, references, and potential implications of results.

STUDY DESIGN
Type of study proposed (e.g. randomized controlled trial, prospective cohort study, etc.). Explains the way data will be accessed and gathered.

NATURE OF INTERVENTION(S)
Description of study products, counseling programs, etc. (if applicable).

STUDY POPULATIONS
Salient characteristics of target study population (key inclusion/exclusion criteria).

STUDY SIZE OR POWER
An estimate of total number of participants needed in the study and number in each arm, if applicable or the power to detect an effect. Request biostatistics help if necessary.

HYPOTHESIS(Optional)
An assumption taken to be true for the purpose of argument or investigation. Expected outcome(s) of the study.

FEASIBILITY
Ability to answer question fully, resources, barriers, timeline, supplies, scheduling, data collection, analyses, etc.

DATA(Optional)
Type of data needed to be collected.

INSTRUMENTS(Optional)
Tools needed to measure data.
What Can a Biostatisticians Do for You?

- **Research Design**
  - The Research Question
  - Feasibility of the Study
  - Sample Size
  - Power
  - Randomization
  - Data Collection and Management

- **Analysis**
  - Data Cleaning
  - Descriptive Statistics
  - Advanced Analysis
  - Results

- **Reporting**
  - Tables and Figures
  - Interpretations
  - Scientific Writing
  - Proof Reading
Characteristics of the Sample
Choosing the Sample

- **Representative of the Population**
- **Inclusion Criteria**
  - Specifies the population relevant to the research question and efficient for the study.
- **Exclusion Criteria**
  - Specifies a subset of the population that will not be studied because it would divert focus from the true condition of interest.
Choosing the Sample: Bias

- Bias related to sample collection is typically systematic error
  - Recruitment Bias
    - Ex: Recruiting people out of the phone book
  - Selection Bias - Differences in the comparison group attributable to incomplete randomization
    - Ex: Consenting patients differ from non-consenting
  - Response Rate – Proportion of eligible participants who participate in the study
  - Convenience Sample – Choosing a readily available sample
    - Ex: The first 50 patients that arrive at Shade Tree Clinic
Group Exercise: Characterizing the Sample

- Question from Part 1: Does inclusion of coordinating nurses minimize emergency service in a chronic healthcare service?
- Do we need inclusion/exclusion criteria?
  - How will we gain access to the study population?
  - How can we ensure a good response rate?
- How should we revise the research question (if at all)?
How Do I Get Statistical Significance? Sample Size

- Sample Size
  - How large should my sample be?
    - PS software by William Dupont and Walton Plummer; compatible with Windows
    - Google: sample size Vanderbilt
      http://biostat.mc.vanderbilt.edu/wiki/Main/PowerSampleSize
  - Should I recruit as many subjects as possible?
How Do I Get Statistical Significance? Power

- **Power**
  - We want to show that the average copay is not $17. We hypothesize that it is $17 and search for evidence to support our claim.
  - For our example, *power* is the probability of correctly finding statistical evidence that the average patient's copay is not $17.
  - Conceptually, *power* is the probability of correctly concluding that statistical evidence exists.
  - Acceptable Power
How Do I Get Statistical Significance? Power

- **Clinical Significance vs. Statistical Significance**

- We simulated a sample of 50,000 men and 50,000 women. The average age of the men and women were 28.9 and 29.5 respectively. A statistical test (t-test) found a statistically significant difference between the two groups ($p$-value < .0001)

- **Biostatisticians ≠ Significant**

- **Biostatistician = Better Chance**
Minimizing Sample Size and Maximizing Power

- Conduct a Pilot Study
- More Precise Measurements
  - Calibrated Tools
  - Better Training
  - Automate the Tool
- Use Outcome with More Frequent Events
- Continuous vs. Categorical Variables
At the end of this session, you will:

- Know multiple data collection methods
- Understand the importance of collecting information on confounding variables
- Define at least 5 key steps in designing a project
- Define at least 5 reasons why papers are rejected
Measures and Outcomes: Tools for Quantitative Measures

- Survey/Questionnaire
- Interview (Qual or Quan)
- Observation
- Lab
- Chart Review
- Experiment
- Simulations
- Rubrics/Tests

Are your methods and/or instruments reliable and valid?
Measures and Outcomes: Measuring Operational Definitions

- How do you measure intelligence?
- All of these variables are difficult to operationalize.
- Critics often disagree about measures.
- It is important to find established, tested instruments.
- Talk to experts.
- Review the literature.
Does inclusion of coordinating nurses minimize emergency service in a chronic healthcare service?

For our question, what would be the best data collection method?

Are there any operational definitions that we need to be concerned with?

Do we need to revise our question?
Measures and Outcomes

- Client: “I am interested in knowing if smoking cigarettes is associated with being diagnosed with lung cancer for teenagers who smoked less than 6 months? It's a small study, so I plan to collect the participants' demographics and their smoking status.”

- Biostatistician: So if I understand you correctly, you don't plan to collect information regarding their environment such as exposure to asbestos, radon, or perhaps marijuana use?

- Client: “Oh no! We already know those to be risk factors. For this study, I only care about the association of cigarette smoking.”
Measures and Outcomes

- By a show of hands, who all believes the amount of ice cream sold in America is associated with the number of murders committed?
Measures and Outcomes:

Relation of Ice Cream Sales to Murders for Cities in America

- Number of Murders/year
- Ice Cream Sales (thousands)
Measures and Outcomes:
Measures and Outcomes: Confounding

Relation of Ice Cream Sales to Murders for Cities in America

Season is a confounder.
Measures and Outcomes: Confounding

- “Seven homicides in New York City. None connected in any way but this: They happened during the summer months, when the temperatures rise, people hit the streets, and New York becomes a more lethal place (New York Times: http://www.nytimes.com/2009/06/19/nyregion/19murder.html).”

- “Ice cream consumption is highest during July and August. July is National Ice Cream Month (http://www.makeicecream.com/contriv.html).”
Measures and Outcomes: Confounding

- Confounding – When there is a third factor associated with the outcome and other factors of interest.

![Diagram showing Season pointing to Ice Cream and Murders]

- The take home message is: if there are confounders, you will want to make sure to collect them; otherwise, your results will have serious limitations!
Measures and Outcomes: Rare Events

- Typical methods may not be appropriate when there are few events.
- Generalizing can prove to be challenging.
- Number of events per factor may limit how much we can model.
Ex: Suppose our outcome is a rare leukemia (Y/N) and there are 42 events. To run advanced statistics using information from potential risk factors and confounders, I may suggest using information from 2-4 risk factors and confounders. And this REALLY depends on what type of measurement was taken, the type of analysis, and how it looks when graphed.
Group Exercise

- Does inclusion of coordinating nurses minimize emergency service in a chronic healthcare service?
- What other variables do we need to collect to study our question?
- Will we have the time to address the research question? The sample size? The resources?
- If the answer is “No” revise your question.
Measures and Outcomes: Validity and Reliability

- **Validity** is a measure of how well the phenomenon is represented.

- **Reliability** – The degree to which a measurement is reproducible.

- There are multiple types of validity and reliability that goes beyond the scope of this workshop.

- Always try to use a validated, reliable measure.

- It's important to pretest or pilot study the measure before conducting research.
Critiquing Research Questions

- During their third year of medical school, how do Vanderbilt medical students describe their empathy towards chronically ill patients who express themselves as being in a lot of pain (Part 1)?
- During their third year of medical school, what are the changes in empathy of Vanderbilt medical students towards chronically ill patients who express themselves as being in a lot of pain?
During their third year of medical school, what are the changes in empathy of Vanderbilt medical students towards chronically ill patients who express themselves as being in a lot of pain?
Key Steps of Designing a Project

- This is a not an exhaustive, linear process.

- **Throughout the entire process**, talk to key people: mentor(s), biostatistician, IRB representative.

- Create a good research question.

- Review the literature for creativity, inspiration, methodology, design, limitations, gaps, etc.

- Calculate the sample size, number of allowable variables, and/or power.
Key Steps of Designing a Project

- Determine methods of measurement.
- Determine the key variables.
- If possible conduct a pilot study or pretest.
- Be careful of missing data.
- Know the type of analysis that will be conducted BEFORE finalizing your research proposal.
- Create a timeline.
- Assure access to the subjects.
- Consider geographical barriers.
Research Proposal Template

STUDY TITLE

COMMITTEE
Your name, mentor, and others working on the project.

RESEARCH QUESTION
Concise, informative question that outlines and guides the methods, population, and other details pertinent to developing the study.

RATIONALE
Include relevant background information, references, and potential implications of results.

STUDY DESIGN
Type of study proposed (e.g. randomized controlled trial, prospective cohort study, etc.). Explains the way data will be accessed and gathered.

NATURE OF INTERVENTION(S)
Description of study products, counseling programs, etc. (if applicable).

STUDY POPULATIONS
Salient characteristics of target study population (key inclusion/exclusion criteria).

STUDY SIZE OR POWER
An estimate of total number of participants needed in the study and number in each arm, if applicable or the power to detect an effect. Request biostatistics help if necessary.

HYPOTHESIS(OPTIONAL)
An assumption taken to be true for the purpose of argument or investigation. Expected outcome(s) of the study.

FEASIBILITY
Ability to answer question fully, resources, barriers, timeline, supplies, scheduling, data collection, analyses, etc.

DATA(OPTIONAL)
Type of data needed to be collected.

INSTRUMENTS(OPTIONAL)
Tools needed to measure data.
Some Available Vanderbilt University Resources

- Daily Biostatistics Clinic (http://biostat.mc.vanderbilt.edu/wiki/Main/Clinics)
- Biostatistics Collaboration Center (http://biostat.mc.vanderbilt.edu/wiki/Main/BCC)
- Biostatistics Collaboration Plan (http://biostat.mc.vanderbilt.edu/wiki/Main/CollaborationAssignments)
- CORE Conversations (https://medschool.vanderbilt.edu/otlm/core/core_conversations)
- CORE Collaborative in Qualitative Research (limited); Contact Regina Russell at regina.russell@Vanderbilt.Edu
- Design Studio and Vanderbilt Institute for Clinical and Translational Research (http://www.mc.vanderbilt.edu/victr/pub/)
Reasons Why Papers are Rejected

- Poor research design
- Poor methods section
- Unsupported conclusions
- Unoriginal research
- Failure to collect key variables - especially confounders

- Sample not representative of the population
- Sample size too small
- Incorrect analysis
- Author's hypothesis not tested
- Poor writing
Acknowledgments and References

Acknowledgements to Cathy Jenkins, Regina Russell, Lillian Nanny, and Katherine Allen for constructive criticism, editing, and ideas.


